
Semantische Patentanalyse mit dem Knowledgist und PIA Grundlagen – Beispiele – Kritik

Lothar Walter, Anja Geritz, Martin G. Möhrle

Universität Bremen

Lehrstuhl für BWL - Innovation und Kompetenztransfer

Abstract

Patente stellen einen unerschöpflichen Ideenpool dar. Doch die zunehmende Flut an Patentdokumenten stellt Patentfachleute, Erfinder und FuE-Mitarbeiter vor immer größer werdende Probleme, sich diesen Ideenpool zu erschließen.

Für eine patentorientierte Ideenfindung bieten computergestützte semantische Textauswertungen eine wertvolle Hilfe. So lässt sich beispielsweise mit der Software Knowledgist™ 2.5 mittels eines semantischen Prozessors in sehr kurzer Zeit eine große Zahl von Patentdokumenten analysieren und relevantes Wissen aus den Dokumenten extrahieren. Sogenannte SAO-Strukturen (Subjekt-Aktion-Objekt) bilden hierbei die Kernaussage der analysierten Dokumente ab. Diese entsprechen einer Problem-Lösungsstruktur innerhalb eines „gelesenen“ Patentdokuments.

Eine weitere, computergestützte statistische Aufbereitung der extrahierten Problem-Lösungsstrukturen mit dem am Lehrstuhl entwickelten Softwaretool PIA (Patent-Information-Analysetool) erlaubt das systematische Erkennen vieler umgesetzter Ideen zur Lösung von Erfindungsaufgaben.

Für die unternehmerische Praxis bietet der Einsatz patentorientierter Ideenfindung innerhalb der FuE-Abteilung eine Möglichkeit der Forcierung von Entwicklungsprozessen und interessante neue Impulse für eine Wettbewerberanalyse im strategischen Management.

Einleitung

Schon Genrich S. Altschuller hat die Patentliteratur als einen unerschöpflichen Ideenpool erkannt¹. Doch die immer schneller ansteigende Flut an explizitem Wissen, wie sie sich durch das kontinuierliche Ansteigen von Patentanmeldungen zeigt, stellt die Entwickler, Wissenschaftler, Patentfachleute und –anwälte vor immer größere Probleme. Gerade unter dem Aspekt immer enger begrenzter Zeitressourcen ist das Lesen aller neu veröffentlichten Patentschriften ohne den Einsatz moderner Computertechnik kaum zu bewältigen. Dieser Flut kann mit semantischen Patentanalysen begegnet werden.

Semantik befasst sich mit der Bedeutungsanalyse sprachlicher Ausdrücke. Die Firma Invention Machine Corporation, mit Sitz in Boston USA, bietet seit geraumer Zeit eine Wissensextraktionssoftware, die sich der Semantik bedient, an. Sie kann mit Hilfe eines semantischen Prozessors den Sinn und die Zusammenhänge innerhalb eines Textes erkennen und relevantes Wissen aus ausgewählten Dokumenten extrahieren. Diese Software bietet somit die Möglichkeit, in den zu analysierenden Dokumenten zu suchen, zu lesen, zusammenzufassen und Inhalte zu strukturieren. Die Kernaussagen der analysierten Dokumente werden in einem Subjekt-Aktion-Objekt-Format (SAO-Format) dargestellt. Die Aktion und das betreffende Objekt stellen das zu behandelnde Problem oder die jeweilige Aufgabenstellung dar, das Subjekt ist die Lösung des Problems. Es ist naheliegend, dass eine solche Aufdeckung einer Problem-Lösungsstruktur in den analysierten Dokumenten durch eine semantische Textanalyse wesentliche Vorteile im Vergleich zu der Schlüsselwortsuche durch die verschiedensten Suchmaschinen bietet. Die Auflistung von SAOs erlaubt eine schnelle Identifikation von Problem-Lösungsstrukturen und damit ein gezieltes Selektieren von relevanten Patentdokumenten. Neben dem Einsatz von Wissensextraktionssoftware zur SAO-Bestimmung ermöglicht eine weitere computergestützte statistische Aufbereitung mit dem Softwaretool PIA (Patent-Information-Analysetool) die Bestimmung von Ähnlichkeiten in den semantischen Strukturen. Ein Vergleich von SAO-Strukturen bzw. SAO-Pools von Unternehmen und/oder Technologiefeldern zur Wettbewerberanalyse ist nun möglich.

Problembezogene Patentrecherchen auf Basis der Semantik erschließen somit ganz neue Wege im Innovations- und Patentmanagement. Sie helfen dem Innovationsmanagement bei der Ideenfindung oder liefern neue Ansätze für Mergers & Aquisitions. Auch bieten sie dem Patentmanagement neue

¹ Altschuller, G. S.: Erfinden – Wege zur Lösung technischer Probleme, 2. Auflage. Cottbus: Planung und Innovation 1998.

Bewertungsansätze mittels SAO-Cluster oder erlauben eine Komprimierung des relevanten Patenwissens in SAO-Strukturen.

Dieser Aufsatz gliedert sich in fünf Abschnitte:

1. Mittels Stichwortsuche können Patentrecherchen in klassischer Weise zur Beantwortung rechts- und informationsbezogener Fragestellungen herangezogen werden. Ein neuer Ansatz ist die semantische Analyse der Patente, die inhaltliche und problembezogene Fragestellungen beantworten kann.
2. Für das Verständnis der Funktionsweise von semantischen Text- und Patentanalysen werden die philosophisch-sprachwissenschaftlichen Begriffe der Semiotik und deren Teildisziplinen, der Syntaktik, der Semantik und der Pragmatik, definiert.
3. Der semantische Prozessor ist das Kernstück der Software Knowledgist™ 2.5. Er extrahiert aus den zu analysierenden Text- und/oder Patentschriften technische Problemlösungsstrukturen, sogenannte SAO-Strukturen, welche die Basis für eine weitere statistische Auswertung mit dem Patent-Informationen-Analysetool (PIA) darstellen und auf der denen patentorientierte Ideenfindung basiert.
4. An einem Fallbeispiel wird die Durchführung einer semantischen Patentanalyse demonstriert, wie sie in der unternehmerischen Praxis zur Ideenfindung angewendet werden kann.

1 Patentrecherchen

Die Patentliteratur stellt weltweit ein einmaliges und umfangreiches technisches Wissensdokument dar. Es wird geschätzt, dass ca. 80% des technischen Wissens in der Patentliteratur und ca. 20% anderweitig veröffentlicht ist². Das in Patentdokumenten niedergeschriebene Wissen ist für die Wirtschaft im allgemeinen und für den Bereich Forschung und Entwicklung im besonderen von unschätzbarem Wert. Durch die verpflichtende Veröffentlichung dieses Wissens in Form von Patentschriften sind kontinuierlich umfassende und aktuelle Informationen über technische Problemlösungen und den Stand der Technik zu erhalten. Die Analyse dieser Patentliteratur birgt jedoch große Probleme in sich, da die Zahl der veröffentlichten Patentdokumente stark wächst. Eine Analyse muss somit gezielt und durch den Einsatz moderner Computertechnik vorgenommen werden. Stehen Fragen hinsichtlich des Schutzzumfanges von Erfindungen im Vordergrund des Interesses, können rechtsbezogene Recherchen durchgeführt werden, bei technischen Fragestellungen kann dagegen informationsbezogen recherchiert werden.

² Barske, H.: Innovations-Vorsprung. Strategische Erneuerung – Produktinnovation – Patentwesen. Symposium Publishing, Düsseldorf 2001.

Rechtsbezogene Patentrecherchen führt man vorwiegend nach Abschluss von Entwicklungsarbeiten durch. Sie sollen möglichst vollständig inhaltsgleiche Lösungen ermitteln und dienen der Feststellung der Rechtsbeständigkeit eines Schutzrechtes. Unter rechtsbezogene Patentrecherchen fallen dabei Patentfamilienrecherchen (In welchen Ländern ist eine Erfindung patentiert?), Namensrecherchen (Welcher Erfinder und welches Unternehmen hat was patentiert?), Verletzungsrecherchen (Existieren Patente, die schon früher einen interessierenden Aspekt behandelt haben und somit neuheitsschädlich sind?), Neuheitsrecherchen (Gibt es Patente, die der eigenen Patentanmeldung entgegenstehen?), Rechtsstandsrecherchen und Recherchen nach Entgegenhaltungen (Welchen rechtlichen Status hat ein bestimmtes Patent? Ist das Patent erteilt, laufen Einspruchsverfahren oder gibt es ein Nichtigkeitsverfahren?). Bei all diesen Recherchearten wird mittels Stichworten in verschiedenen Patentdatenbanken nach den entsprechenden Dokumenten gesucht.

Informationsbezogene Recherchen werden im allgemeinen im Anfangsstadium von Entwicklungsarbeiten durchgeführt. Man unterscheidet zwischen einer Sachrecherche und einer Überwachungsrecherche. Eine Sachrecherche dient zur Unterstützung von Neuheitsrecherchen und Verletzungsrecherchen und wird breit angelegt. Eine Überwachungsrecherche hat schutzrechtliche Motive und dient der technologischen Beobachtung des Wettbewerbs. Es gilt zu beantworten, zu welchen Patentgruppen, Unternehmens- oder Erfindernamen neue Patentanmeldungen getätigt worden sind. Bei diesen informationsbezogenen Recherchen wird ebenfalls mittels Stichworten in den Patentdokumenten recherchiert. Sie sollen eine übersichtliche Zusammenstellung der technischen Lösungen und Entwicklungslinien zu einer offenen Problemstellung ermitteln und dienen der Vermeidung von Doppelentwicklungen und Schutzrechtskollisionen.

Diese in klassischer Weise durchzuführenden Patentrecherchen sind hilfreich, um den Stand der Technik zu ermitteln, Unternehmensanalysen zu erstellen, technologische Durchbrüche aufzuspüren, Kooperationspartner zu finden und Anregungen zu gewinnen. Besteht jedoch Interesse an inhaltlichen und problembezogenen Fragestellungen, so kann eine Antwort nur mittels einer modernen semantischen Textanalyse in den Patentdokumenten gegeben werden. Bevor auf die semantische Patentanalyse eingegangen und ein Fallbeispiel beschrieben wird, soll für das weitere Verständnis nachfolgend erläutert werden, was unter dem Begriff Semiotik zu verstehen ist.

2 Semiotik: Syntaktik – Semantik - Pragmatik

Die menschliche Sprache wird schon seit Jahrhunderten intensiv untersucht. Versuche, Sprache zu deuten, gehen in die Anfänge von Philosophie und Wissenschaft zurück. Bereits Platon versuchte mit seinem sprachphilosophischen Text „Kratylos“ eine Deutung von Sprache und ihrer Zusammenhänge vorzunehmen³. Heute bereichert die Linguistik, wie dieser Wissenschaftszweig benannt wird, viele Disziplinen, wie zum Beispiel die Kontaktlinguistik, die Psycholinguistik, die Computational Linguistics (CL) oder das Natural Language Processing (NLP). Die semantische Patentanalyse, die aus den beiden obengenannten Teildisziplinen der Informatik hervorgeht, hat ihr theoretisches Fundament ebenfalls in der Linguistik. Die Linguistik allgemein beinhaltet auch die Disziplin der Semiotik mit den Teilgebieten der Syntax oder Syntaktik, der Semantik und der Pragmatik.

Unter Semiotik wird seit Charles S. Peirces Abhandlung „On a New List of Categories“ die Lehre von den am Sprachprozess beteiligten Zeichen bzw. Zeichenreihen oder die Lehre der Kennzeichen verstanden^{4,5,6}. Durch Charles Morris Arbeit „Foundations of the Theory of Signs“ erfuhr die Semiotik eine weitere Unterteilung in die Teildisziplinen Syntax oder Syntaktik, Semantik und Pragmatik. Der Zusammenhang zwischen diesen drei Teildisziplinen wird durch das sogenannte *semiotische Dreieck* verdeutlicht, welches auf den Arbeiten von Charles K. Ogden und Ivor A. Richards fußt⁷. In Verbindung mit der durch Morris befürworteten Unterteilung der Semiotik in Syntax/Syntaktik, Semantik und Pragmatik kann das *semiotische Dreieck* dementsprechend umgedeutet werden (siehe Abbildung 1)⁶.

Jedes Zeichen weist demnach bestimmte Relationen auf. Nach Morris werden sie als Dimensionen, die nach wachsender Komplexität geordnet sind, aufgefasst.

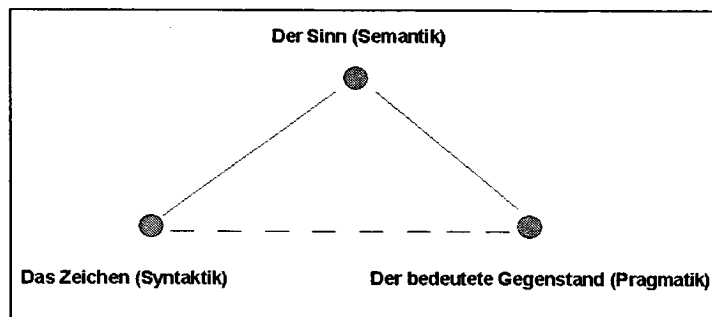


Abbildung 1: Das semiotische Dreieck (in Anlehnung an [6]).

³ vgl. Mittelstraß, J.: (Hrsg.) Enzyklopädie Philosophie und Wissenschaftstheorie. Bibliographisches Institut: Mannheim, Wien, Zürich 1984.

⁴ vgl. Hörz, H.; Liebscher, H.; Löther, R.; Schmutzer, E.; Wollgast, S.: Philosophie und Naturwissenschaften – Wörterbuch zu den philosophischen Fragen der Naturwissenschaften. Sonderausgabe 3., vollständig überarbeitete Auflage, Wiesbaden 1996

⁵ vgl. Prechtl, P.; Burkhard, F.-P.: (Hrsg.) Metzler-Philosophie-Lexikon – Begriffe und Definitionen. Stuttgart, Weimar 1996.

⁶ vgl. Glück, H.: (Hrsg.) Metzler-Lexikon Sprache. Metzler Verlag, Stuttgart, Weimar 2000.

⁷ Bußmann, H.: Lexikon der Sprachwissenschaften. 2.Aufl. Kröner, Stuttgart 1990.

Hierbei handelt es sich⁵:

1. um die syntaktische Dimension. Sie verdeutlicht die Relation zwischen den Zeichen (im *semiotischen Dreieck* als *andere Zeichenträger* benannt).
2. um die semantische Dimension. Sie bezeichnet die Relation zwischen Zeichen und Bedeutung (im *semiotischen Dreieck* als *Designat* oder *Denotat* benannt) und
3. um die pragmatische Dimension. Sie bezeichnet die Relation zwischen Zeichen und Zeichenbenutzer (im *semiotischen Dreieck* als *Interpretant* oder *Interpret* benannt).

Unter *Syntaktik* wird allgemein die sprachwissenschaftliche Untersuchung der internen Struktur von Zeichensystemen, d.h. die Beziehung der Zeichen untereinander, verstanden^{3,8,9}. Jede Sprache besteht aus einem System von Regeln, die es ermöglichen aus dem Zeichenvorrat gültige Ausdrücke oder Sätze zu formulieren. Jede Syntax einer Sprache wird als formales System verstanden. Natürliche Sprachen sind im Vergleich zu künstlichen Sprachen, wie sie zum Beispiel in der Computerprogrammierung verwendet werden, weitaus komplexer aufgebaut⁴. Die Syntax oder Syntaktik bildet die Grundlage für die Semantik.

Der Begriff *Semantik* wird in den Sprachwissenschaften nicht immer einheitlich verwendet⁴. Seit Michel Bréals „*Essai de sémantique, science des significations*“ wird die *Semantik* als eine Teildisziplin der Linguistik aufgefasst. Diese Bezeichnung löste den damals gängigen Begriff der Semasiologie ab³. Im weiteren Verlauf soll unter dem Begriff *Semantik* die Lehre von den sprachlichen Zeichen und Zeichenfolgen oder auch die Bedeutung sprachlicher Begriffe subsumiert werden⁷. Auf die Zuordnung neuerer Arbeiten auf diesem Gebiet zu einzelnen Bedeutungstheorien wird im Rahmen dieses Aufsatzes bewusst verzichtet¹⁰. Die *Semantik* ist der Ausgangspunkt für die Teildisziplin der *Pragmatik*.

Der dritte und letzte Begriff, der im Rahmen des semiotischen Dreiecks aufgegriffen wird, wird als *Pragmatik* bezeichnet. Er stellt den komplexesten Teil der Semiotik dar. Die Teildisziplin der *Pragmatik* untersucht die Relation zwischen Zeichen und Benutzer⁸. Gegenstand der Untersuchung sind die Absichten, Meinungen und Handlungen, die Benutzer von Sprache haben⁹. Die

⁸ Ulrich, W.: Wörterbuch Linguistische Grundbegriffe. 5.Aufl., Gebrüder Borntraeger Verlagsbuchhandlung, Berlin, Stuttgart 2002.

⁹ Hügli, A.; Lübcke, P. (Hrsg.): Philosophielexikon. Personen und Begriffe der abendländischen Philosophie von der Antike bis zur Gegenwart. Vollst. überarb. und erw. Neuaufl. Reinbek bei Hamburg: Rowohlt 1997.

¹⁰ von Stechow, A.; Wunderlich, D.: Semantik. Ein internationales Handbuch der zeitgenössischen Forschung = Semantics. Handbücher zur Sprach- und Kommunikationswissenschaft. Bd. 6 de Gruyter, Berlin, New York 1991.

Pragmatik übernimmt unterschiedliche Funktionen, wie zum Beispiel eine Symbolfunktion, eine Symptomfunktion, eine Signalfunktion und eine Bewertungsfunktion. Unter der Symbolfunktion wird die Beschreibung von Ereignissen und Sachverhalten mittels Sprache verstanden. Durch die Symptomfunktion können Gefühle ausgedrückt werden. Die Signalfunktion erzeugt ein spezielles Verhalten, das durch den Benutzer von Sprache erwünscht ist. Die Bewertungsfunktion soll Denkweisen oder Zustände mit bestimmten Prädikaten belegen ([4] und Problem-Lösungsstruktur beim Knowledgegist™ 2.5).

Zusammenfassend wird deutlich, dass die Semiotik ein Grundwerkzeug für das Verständnis von Sprache ist. IT-gestützte Software zur Verarbeitung von Sprache muss sich diese Grundlagen zu eigen machen, um Sprache verstehen und verarbeiten zu können. Im nächsten Abschnitt soll die konkrete Umsetzung der linguistischen Grundlagen in Form des semantischen Prozessors erläutert werden. Dabei wird die abgestufte Vorgehensweise des in der Software Knowledgegist™ 2.5 implementierten semantischen Prozessors verdeutlicht.

3 Der semantische Prozessor

Täglich werden etwa 5.000 Patente angemeldet. Das bedeutet, alle 20 Sekunden reicht ein Erfinder bei irgendeinem Patentamt, irgendwo auf dieser Welt eine Patentanmeldung ein¹¹. Die für ein Unternehmen interessanten Wissensmengen wachsen kontinuierlich an.

Wie kann der Wissensflut begegnet werden? Und wie wird dabei vorgegangen?

Die im vorangegangenen Abschnitt erläuterten sprachwissenschaftlichen Grundlagen bilden die Basis der semantischen Patentanalyse. Eine Lösung zur Beherrschung der Wissensflut ist die Umsetzung dieser Theorien in IT-unterstützte Software. Dies wird als Natural Language Processing (NLP), Linguistic Analysis (LA) oder Computational Linguistics (CL) bezeichnet. In den letzten zweieinhalb Jahrzehnten erfreute sich dieses Forschungsgebiet zunehmender Bedeutung. Von 1975 bis zum heutigen Tag wurden insgesamt 419 Patente beim United States Patent and Trademark Office (USPTO) angemeldet, die sich mit dem Themenkreis Natural Language Processing (NLP) befassen (eigene Recherchen – Stand 01.08.2002). Seit über zwanzig Jahren wird bereits mit Hochdruck an IT-Lösungen gearbeitet, die es ermöglichen, Texte mit dem Computer zu lesen und zu verstehen. Erst mit der drastischen Steigerung der Rechnerleistungen im Laufe des letzten Jahrzehnts sind viele dieser Anwendungen möglich geworden¹².

¹¹ Lüthy, W.; Schweizer, P.: Sag mir, was in den Dokumenten steht! In: TRIZ.ch – erschienen im World Wide Web (URL: <http://www.triz.ch/KNPaper.pdf> (06.03.2002))

¹² Invention Machine: Knowledgegist – Broschüre. <http://www.invention-machine.com/prodserv/knowledgegist.cfm> 2002

Eine kommerziell angebotene Lösung zur Beherrschung und Verarbeitung u.a. großer Patentdatenmengen ist der Knowledgist™ 2.5¹³. Ein Werkzeug, das sich des Natural Language Processings (NLP) bedient. Es beinhaltet eine Vielzahl unterschiedlicher Hilfswerkzeuge, die den Analyseprozess von Patenten unterstützt.

Nachfolgend soll der Ablauf einer semantischen Patentanalyse illustriert und jeder vom semantischen Prozessor durchgeführte Schritt erläutert werden. Wie in Abbildung 2 gezeigt, sind vom Patent bis zur SAO-Struktur fünf Analyseschritte zu durchlaufen.

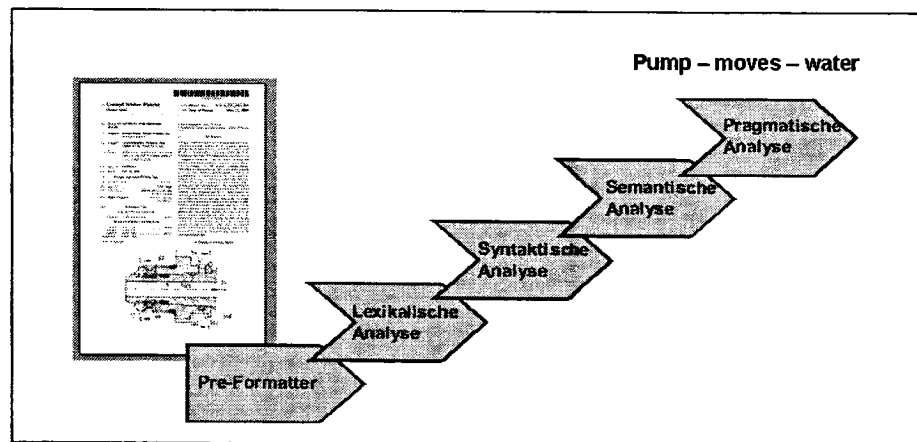


Abbildung 2: Abfolge der Analyseschritte vom Patent zur SAO-Struktur

Ausgangspunkt der semantischen Patentanalyse ist das Patentdokument, welches durch die Abarbeitung von fünf Analyseschritten in eine SAO-Struktur umgewandelt wird. Die SAO-Struktur (engl.: Subject-Action-Object) spiegelt eine technische Ursache-Wirkungsbeziehung wider, die in jeder Patentschrift enthalten ist. Zur Zeit ist eine semantische Patentanalyse nur in englischer Sprache möglich (vgl. hierzu und im folgenden [14]).

- (1) Im ersten Schritt, der *Pre-Formatter-Analyse*, wird das Patentdokument vorformatiert. Die Patentschrift wird in Textformat umgewandelt, Abbildungen und andere nichttextliche Elemente werden entfernt. Vorhandene lexikalische Fehler werden mit Hilfe eines komplexen Wörterbuchs korrigiert. Eine Gliederung des Patenttextes in einzelne Sätze wird vorgenommen.
- (2) Im zweiten Schritt unterzieht der semantische Prozessor die einzelnen Sätze einer *lexikalischen Analyse*. Dies geschieht mit Hilfe eines Wörterbuchs, welches Informationen über denkbare Wortklassen der jeweils betrachteten Satzelemente enthält. Jedes Wort wird mit mindestens einem Bezeichner

¹³ Stauch, R.: Manuskript über Knowledgist. Invention Machine germany GmbH, Ingolstadt, Manuskript 2001.

¹⁴ Invention Machine: White Paper – Accelerate your speed to Knowledge. Invention Machine Corp., p.1-21, 2000

versehen. Dieser Bezeichner steht für eine Wortklasse, wie z.B. Verb, Subjekt, Adjektiv oder Präposition.

- (3) Im dritten Schritt wird eine *syntaktische Analyse* der einzelnen bezeichneten Wörter vorgenommen. Diese Analyse ist in zwei Zwischenschritte unterteilt. Wie im vorangegangenen Abschnitt deutlich wurde, dient die Syntax dazu herauszufinden, welche Struktur ein gültiger Satz annehmen kann. (i) Im ersten Zwischenschritt werden daher die vorhandenen syntaktischen Regeln angewandt. Dies führt zu einer Reduktion der möglichen Bezeichner. (ii) Im zweiten Zwischenschritt werden statistische Auswertungen herangezogen, die Aufschluss darüber geben, welche Reihung von Wortklassen überhaupt logisch sind. Hierzu wurden im Vorfeld der Programmierung des semantischen Prozessors technische Dokumente aller Art und im besonderen Maße Patente analysiert und entsprechende Reihungswahrscheinlichkeiten berechnet. Als Unterstützung wird das Hidden-Markov-Modell (HMM) hinzugezogen¹⁵. Mit diesem Modell sind die Wortklassen, die mit großer Wahrscheinlichkeit vorliegen, determinierbar.
- (4) Im Rahmen der semantischen Analyse wird die *Bedeutung der Zeichen* näher untersucht⁸. Die Bedeutung der Satzelemente im Satzzusammenhang soll geklärt werden. Hierzu sind die Hauptstrukturelemente Subjekt (engl.: Subject), Prädikat (engl.: Action) und Objekt (engl.: Object) zu identifizieren. Eine sinnvolle Relation dieser Elemente wird im Rahmen der semantischen Analyse hergestellt. Durch die SAO-Relation kann der Sinn eines Satzes ausreichend gut wiedergegeben werden.
- (5) Als letzten Schritt der Patentanalyse führt der semantische Prozessor eine *pragmatische Analyse* durch. Das bedeutet, die Beziehung zwischen den Zeichen, in diesem Fall der Satzelemente und des Benutzers, sollen herausgestellt werden. Das Ergebnis ist die Identifikation der im Patent enthaltenen Problem-Lösungsstrukturen. Die bereits identifizierte SAO-Struktur wird daher als ein Problem (Action-Object repräsentieren das Problem) und seine Lösung (Subject repräsentiert die Problemlösung) interpretiert.

Im Rahmen dieses Abschnitts wurde grob die Vorgehensweise des im KnowledgeGist™ 2.5 implementierten Prozessors vorgestellt. Diese Art der Patentanalyse bietet eine ausbaufähige Grundlage für weitere Patentanalyseansätze an, wie der patentorientierten Ideenfindung.

¹⁵ Rabiner, L. R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE, Vol. 77, Nr. 2, S.257-286, 1989.

4 Patentorientierte Ideenfindung

Bei der Durchführung von semantischen Patentanalysen ist neben dem Einsatz des Knowledgist™ 2.5, der die Patenttexte in SAO-Strukturen umwandelt und eine Datenbank extrahierter SAO-Strukturen generiert, der Einsatz eines weiteren Softwaretools notwendig. Um die Ähnlichkeit mehrerer SAO-Strukturen oder zusammengestellter SAO-Pools bestimmen zu können, wurde eigens das Patent-Informationen-Analysetool (PIA) entwickelt. Dieses Tool erlaubt den Vergleich von SAO-Strukturen unterschiedlicher Patentpools und bestimmt deren Ähnlichkeit.

Das Patent-Informationen-Analysetool (PIA) bietet eine Schnittstelle, die es ermöglicht, die durch den Knowledgist™ 2.5 analysierten Patentdokumente und extrahierten Übersichtslisten der SAO-Strukturen einzulesen. Es können wahlweise mehrere Übersichtslisten gemeinsam eingelesen werden, um so den gewünschten SAO-Pool zu erzeugen. Zu den eingelesenen SAO-Strukturen werden zusätzlich weitere bibliographische Daten, wie z.B. IPC-Klassifizierung, Veröffentlichungsdatum oder Anmelder hinzugefügt. Die gesamten Informationen, d.h. SAO-Strukturen und bibliographische Zusatzinformationen, werden in einer Datenbank abgespeichert und sind jederzeit abrufbar.

Um eine semantische Patentanalyse zur patentorientierten Ideenfindung durchzuführen, stehen mit der PIA vier Analysemöglichkeiten zur Verfügung:

1. Durch die *Einzelanalyse* wird ermöglicht, ein Patent aus dem Patentpool auszuwählen und seine Ähnlichkeit mit dem Rest der Patente des Patentpools zu bestimmen.
2. Im Gegensatz dazu steht die *Gesamtanalyse*. Die PIA vergleicht bei der Gesamtanalyse alle Patente des Patentpools miteinander.
3. Eine dritte Möglichkeit stellt die *Analyse nach IPC-Klassifizierung* dar. Der Nutzer terminiert im Vorfeld die für seine Analyse interessante IPC-Klassifizierung. Der PIA wird über ein Auswahlfeld mitgeteilt, welche IPC-Klasse näher untersucht werden soll. Alle Patente der betreffenden IPC-Klasse werden automatisch selektiert und im nächsten Schritt miteinander verglichen. Am Ende der Analyse steht die Berechnung der spezifischen Ähnlichkeiten mittels einer je nach Analysewunsch zu modifizierenden Ähnlichkeitsformel.
4. Die vierte Möglichkeit besteht im *Vergleich zweier Patentgruppen*. Über ein Eingabefeld definiert der Nutzer zwei Patentgruppen. Da jedes Patent in der Datenbank unter einer Identifikationsnummer abgelegt ist, genügt die Eingabe der zwei von der PIA zu betrachtenden Bereiche innerhalb des Patentpools.

Die durch die vier Analysemöglichkeiten mit PIA berechneten Ähnlichkeitswerte werden ebenso in der Datenbank zur weiteren Auswertung abgelegt. Hier bietet PIA dem Nutzer drei Auswertungsmöglichkeiten an:

1. Als erstes kann eine *Auswertung nach Treffertypen* durchgeführt werden. Sie hilft dem Nutzer, Verständnis über die Struktur der Ähnlichkeiten zu erlangen. Der Ähnlichkeitswert wird numerisch dargestellt und gibt keine Aussage über die Art der einzelnen Treffer. Innerhalb der Auswertung nach Treffertypen erfolgt eine detaillierte Auflistung der Treffer nach SAO-, AO- und S-Treffern. Zusätzlich zur Anzeige der Treffer werden bibliographische Daten zu den sich ähnelnden Patenten angeboten.
2. Eine Auswertung *mittels graphischer Darstellung* der Ähnlichkeitswerte in Form eines Histogramms verhilft zu einer schnellen Übersicht über die unterschiedlichen Ähnlichkeitswerte der verglichenen Patente. Diese Darstellung ist eine geeignete Grundlage für weitere Analysen.
3. Die Ähnlichkeitswerte können durch die PIA zur *Auswertung in Form einer Matrix* dargestellt werden. Diese Funktion dient ausschließlich der Vorstufe einer weitergehenden statistischen Analyse der Ähnlichkeitswerte durch zusätzliche Software wie z.B. mit der Statistiksoftware SPSS.

Als Beispiel zur Nutzung semantischer Patentanalysen mit dem Knowldgist™ 2.5 und der PIA sei die patentorientierte Ideenfindung nachfolgend detailliert erläutert. Diese gibt den Entwicklern, Forschern und anderen Interessierten in den FuE-Abteilungen der Unternehmen ein Werkzeug in die Hand, welches es ihnen ermöglicht, Patente aus einer sehr großen, unüberschaubaren Zahl von Patentschriften zu identifizieren, die neue Ideen und Lösungsansätze für das interessierende Unternehmen darstellen. Die patentorientierte Ideenfindung wird in drei Schrittfolgen A, B und C durchgeführt (siehe Abbildung 3).

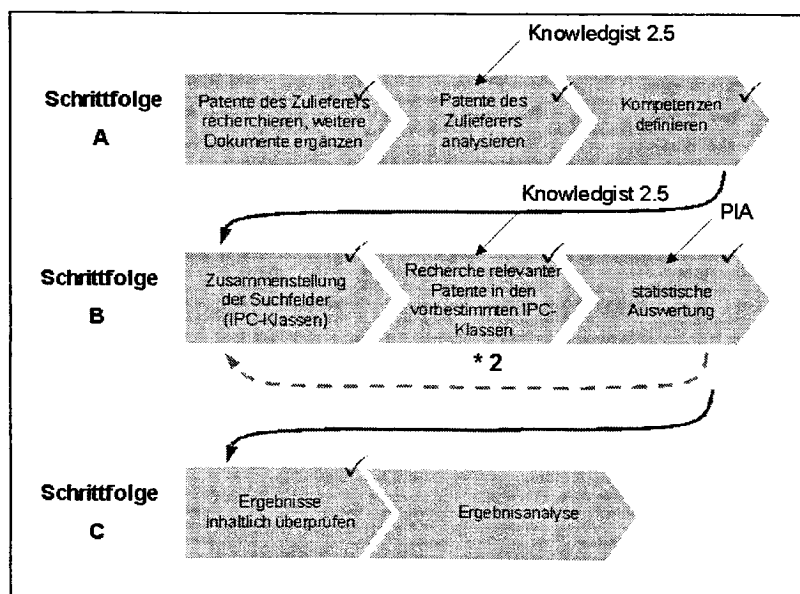


Abbildung 3: Die drei Schrittfolgen A, B und C zur patentorientierten Ideenfindung.

In der *Schrittfolge A* werden die Patente des interessierenden Unternehmens recherchiert und zusammengestellt. Die Extraktion der SAO-Strukturen aus diesen Unternehmenspatenten erfolgt mit dem Knowledgist™ 2.5. Dieser SAO-Pool bildet die technologische Kompetenz des Unternehmens ab.

In der *Schrittfolge B* wird mit dem Unternehmen das zu untersuchende Technologiefeld anhand von IPC-Klassen, Gruppen und Untergruppen eingegrenzt. Danach erfolgt die Recherche aller relevanter Patente in den vorbestimmten IPC-Klassen und wiederum die Extraktion der SAO-Strukturen aus diesen ausgewählten Patenten. Die Anzahl der hier zu analysierenden Patente kann dabei sehr groß sein. Eine sich anschließende statistische Auswertung der beiden SAO-Pools (Unternehmen und Technologiefeld) mit PIA bestimmt deren Ähnlichkeit.

In der *Schrittfolge C* werden die Ergebnisse ausgewertet, indem die Patente, welche die größten SAO-Übereinstimmungen zeigen, zur weiteren Bearbeitung herausgesucht werden. Diese, durch die Ähnlichkeitsanalyse mit PIA ausgewählten Patente beschreiben Erfindungen die zur Ideenfindung herangezogen werden können. Da die Anzahl dieser Patente sehr viel geringer ist, als die Anzahl der Patente aus den vorbestimmten IPC-Klassen, ist ein Lesen und Verstehen manuell in viel kürzerer Zeit möglich. Im nächsten Abschnitt soll konkret eine semantische Patentanalyse, durchgeführt bei einem Automobilzulieferer, zur patentorientierten Ideenfindung vorgestellt werden.

5 Fallbeispiel

Viele national und international agierende Unternehmen bedienen sich der unterschiedlichen Patentanalyse und -recherchearten, um die Arbeit in den FuE-Abteilungen durch bereits in Patenten veröffentlichten Problemlösungen zu bereichern. Dies ist ein legitimes Mittel zur gezielten Stimulation der FuE-Tätigkeit von außen. Das Patent wird durch diese Unternehmen nicht ausschließlich als ein Schutzrecht antizipiert, sondern auch als eine Quelle für neue Ideen. Diese Ideen und Problemlösungen sind im besonderen Maße Translationsleistungen. Problemlösungen aus unternehmensfremden Technologiefeldern sind durchaus geeignet, als Ideenquelle zu dienen. Das zentrale Problem ist jedoch die Identifikation der als Ideenquelle tauglichen Patente. Wie identifiziert man Patente, die Impulse, Ideen und Lösungsansätze anbieten?

Eine patentorientierte Ideenfindung wurde beispielhaft bei einem nationalen Automobilzulieferer auf Basis einer semantischen Patentanalyse mit dem Knowledgist™ 2.5 und der PIA eingesetzt. Dabei wurden die drei empfohlenen Schrittfolgen A, B und C durchlaufen.

In der *Schrittfolge A* erfolgte die Selektion der Patente und die Bildung von Patentpools. Diese Selektion und Bildung der Patentpools ist zwingend not-

wendig, um eine passende und verwendbare Basis für die semantische Patentanalyse zu schaffen. In einem ersten Schritt wurden alle Patente des Automobilzulieferers identifiziert und um ergänzende Dokumentationen und zusätzliches Informationsmaterial erweitert. Hierbei lag der Fokus besonders auf dem Technologiefeld, welches mit Ideen für die FuE gespeist werden sollte. Aus dem Patentpool extrahierte die Software Knowledgist™ 2.5 die enthaltenen SAO-Strukturen. Die SAO-Strukturen sind in einem Workshop einem abteilungsübergreifenden Team zur Bewertung vorgelegt worden. Die Aufgabe der Workshopteilnehmer bestand darin, die den Quellpatenten zugeordneten SAO-Strukturen auf ihre Stichhaltigkeit und Bedeutung zu überprüfen. Im Anschluss an die Bewertung wurden die SAO-Strukturen inhaltlich durch Synonyme erweitert. Dann wurde eine inhaltliche Clusterung der SAO- bzw. AO-Strukturen vorgenommen. Es entstanden objektorientierte Cluster, die die Kernkompetenzen des Unternehmens ausreichend abbildeten.

Die Teilnehmer des Workshops wurden gebeten, die gebildeten Cluster kritisch zu prüfen und gegebenenfalls um zusätzliche Strukturen zu ergänzen. Mit der Definition der Kernkompetenzen des beteiligten Unternehmens endete der Schritt der Selektion des Patentpools.

In der *Schrittfolge B* war die Definition der zu betrachtenden Suchfelder durch das Unternehmen erforderlich. Diese Suchfelder orientierten sich an der IPC-Klassifizierung und bildeten die vom Unternehmen überwachten Technologiefelder ab. Nach der Festlegung dieser Suchfelder wurde in den festgelegten Bereichen eine Patentrecherche vorgenommen. Insgesamt konnten auf diese Weise 19.974 Patente in den einzelnen Suchfeldern lokalisiert werden. Es erfolgte eine Analyse des Patentpools mit dem Knowledgist™ 2.5. Im Anschluss daran wurde eine Gesamtanalyse mit der PIA durchgeführt, um die Patente mit den größten Ähnlichkeiten zum Patentpool des involvierten Unternehmens zu ermitteln. Durch dieses Vorgehen wurden die Patente in den Suchfeldern bestimmt, die ähnliche Problem-Lösungsstrukturen beschreiben wie die Patente aus dem Patentpool des Unternehmens. Als Ergebnis der Gesamtanalyse konnten 182 Patente aus den Suchfeldern entnommen werden.

In *Schrittfolge C* wurden die Patentschriften in vollständiger Textfassung und mit angefügten Zeichnungen ausgewertet. Um die generierten Ergebnisse gezielt zu explorieren, wurde ein zweistufiges Bewertungsverfahren gewählt. Zwei Mitgliedern des Workshopteams wurden die Patentzusammenfassungen sowie alle relevanten Zeichnungen der 182 Patente vorgelegt. In einer ersten Bewertungsrunde evaluierte das Team die Patente unter Zuhilfenahme der ABC-Analyse. Für das Unternehmen besonders wichtige Patente, die als Ideenquelle dienen könnten, wurden mit einem A bewertet. Wichtige Patente, die als Ideenquelle relevante Aspekte enthielten, wurden mit einem B und unwichtige,

irrelevante Patente mit einem C versehen. Ergebnis dieser Bewertungsrunde waren 36 Patente, die mit dem Urteil A bewertet wurden.

In einer zweiten Bewertungsrunde wurde die verbleibende Anzahl der A-Patente mit der PMI-Methode begutachtet. Hierzu sollten die Patente in drei Kategorien, positiv, negativ und interessant, eingeordnet werden. Ergebnis dieses Evaluations-schrittes waren 14 Patente, die als positiv oder interessant in Bezug auf die beschriebene Problemlösung erachtet wurden und generell als Ideenquelle geeignet erschienen. Von diesen 14 Patenten sind 3 Patente in IPC-Klassen kategorisiert, welche nicht vom Unternehmen überwacht werden und nicht als für das Unternehmen wichtig erachtet werden.

Die praktische Anwendung der patentorientierten Ideenfindung erlaubt einen Ausblick auf denkbare Erweiterungen des beschriebenen Verfahrens.

AUSBLICK

Es wurde gezeigt, dass semantische Patentanalysen unter Einsatz des Knowledgist™ 2.5 und der PIA ganz neue Wege im Innovationsmanagement und im Patentmanagement erschließen und für das strategische Management viele neue Impulse liefern.

So erlaubt eine patentorientierte Ideenfindung innerhalb des Innovationsmanagements beispielsweise die Bereitstellung interessanter Patentschriften als Ideenquelle für das FuE-Management. Weiterhin können Ähnlichkeitsanalysen von SAO-Strukturen von Patenten aus verschiedenen Technologiefeldern und/oder Unternehmen zur Wettbewerberanalyse herangezogen werden. Ein neuer prozessorientierter Ansatz für Mergers & Aquisitions ist auf der Grundlage der semantischen Patentanalyse ebenso denkbar. Durch Identifikation von gemeinsamen Problem-Lösungsstrukturen zweier Unternehmen ist eine technologische Verflechtung visualisierbar. Hieraus sind diverse strategische Empfehlungen abzuleiten.

Innerhalb des Patentmanagements kann eine Komprimierung des Patentwissens auf SAO-Strukturen ein Hilfsmittel zum schnellen Finden von Problemlösungen darstellen oder die Darstellung von SAO-Clustern begründen neue Ansätze bei einer Patentbewertung. Somit bieten semantische Patentanalysen und deren statistische Auswertung in Zukunft für die unternehmerische Praxis viele neue Impulse.